

# Effect of Various Spatial Auditory Cues on the Perception of Threat in a First-Person Shooter Video Game

Konstantin Semionov  
contact@konstantinsemionov.com  
Edinburgh Napier University  
Edinburgh, Scotland

Dr Iain McGregor  
i.mcgregor@napier.ac.uk  
Edinburgh Napier University  
Edinburgh, Scotland

## ABSTRACT

This study interviewed game audio professionals to establish the implementation requirements for an experiment to ascertain the effect of different spatial audio localisation systems on the perception of threat in a first-person shooter. In addition, a listening study was carried out involving 35 members of the public, and using three scenes made in Unreal Engine 4 with custom designed sound, authored in Wwise. This established that spatial audio does, indeed, have a noticeable effect on players' perception of threat. Each spatial audio system, however, had different effects on the perception of threat, stealth, realism and position estimation, on all three different visual scenes. Meanwhile, the absence of spatial audio can confuse people and contribute towards inaccurate enemy localisation. With this in mind, a tailored approach to the game design requirements for each project is recommended. Rather than a single, spatialised design for the entire game, each scene should have its own design solution.

## CCS CONCEPTS

• **Applied computing** → **Sound and music computing**; • **Human-centered computing** → **Human computer interaction (HCI)**.

## KEYWORDS

Sound design, first-person shooter, threat perception, audio implementation, spatial audio

### ACM Reference Format:

Konstantin Semionov and Dr Iain McGregor. 2020. Effect of Various Spatial Auditory Cues on the Perception of Threat in a First-Person Shooter Video Game. In *Proceedings of Audio Mostly (AM'20), September 15–17, 2020, Graz, Austria*. ACM, New York, NY, USA, 8 pages. <https://doi.org/10.1145/3411109.3411119>

## 1 INTRODUCTION

Game developers are implementing more advanced audio options in their games [18]. However, there is not enough available research to suggest whether those options enhance gameplay, and if they cause players to experience the game differently. We cannot, using existing research, deduce to what extent do spatial localisation

techniques bring benefit to players of FPS games, and if they make games feel more realistic.

Linking a character's generated sound to an expectation of threat, without the threat being visible, is one sound design technique that can alert a player to a potential threat, and prompt caution in subsequent gameplay [16]. Audio spatialisation techniques are said to provide a way to identify opponent location accurately [42]. Discovering whether these contribute towards the perception of threat within a competitive game would be helpful – as could identifying which are most effective. Localisation and reverberation have been identified as crucial aspects of a first-person shooter experience when enemies are present in the game, due to the “hunter and hunted” nature of the genre [24]. The role of occlusion in spatial localisation, however, is less clear.

Spatial localisation could also be achieved with the use of loudspeaker-based surround systems. Two separate experiments, however, found no significant improvements going from stereo to surround systems in a third-person computer game on either speakers or headphones [42, 43]. Binaural reproduction has been found to convey better spatial attributes compared to stereo for headphone users [51].

Sound source position in the three-dimensional space is determined for a listener using horizontal plane azimuth, elevation in the vertical plane and distance depth. Horizontally, the azimuth is calculated with information arriving at both ears, using interaural time differences (ITD) and interaural intensity differences (IID) [44]. Interaural time difference (ITD) is a human mechanism that localises the source of sound on a horizontal plane, based on the difference between the nearest ear and the furthest ear away from the sound source [35]. Studies have found this system is able to discriminate between angles down to 2° in the 20Hz to 2kHz range [10]. Similar to ITD, interaural intensity difference (IID) is a mechanism that helps with the identification of a sound source position, however, it takes into account difference in sound source volume (intensity) between the ears [35]. Compared to ITD, IID provides better localisation for frequencies above 3 kHz [44].

Vertical localisation is a result of absorption, reflection and diffraction effects, caused by the pinna of the ear as well as the head, shoulders and chest [35, 44]. These spectral cues are called head-related transfer functions or HRTFs and can vary greatly between individual people due to differences in the shape and size of the outer ear [44]. HRTF-based vertical sound localisation has been found to function monaurally in humans for high frequency (more than 7 kHz) content [14]. HRTF-based 3D audio renderers are used in video games to deliver a mixdown of various sound sources within the game. However, it has been suggested that more

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

AM'20, September 15–17, 2020, Graz, Austria

© 2020 Association for Computing Machinery.

ACM ISBN 978-1-4503-7563-4/20/09...\$15.00

<https://doi.org/10.1145/3411109.3411119>

customised HRTF measurements are needed to deliver similar experiences for all players [61]. In extreme cases, non-individualised HRTFs could harm the accurate perception of sound source position [42]. Direct-to-reverberant ratio of sound paired with sound levels are the main contributors towards distance localisation.

### 1.1 3D Audio in Video Games

Generic implementations have succeeded in left-right source positioning in both headphones and speakers, based on IID and ITD. However, in order to achieve better front-back separation, additional head-tracking devices are required for the use with the player's playback device [35]. Absence of head-tracking devices within commercial gaming headphones has meant that filtering is used to simulate front-back positioning instead, which has the drawback of appearing inside the head, rather than behind [35]. This left-right source positioning system is widely used by both Unreal Engine's [23] own audio engine as well as Wwise [3]. HRTF use within video games can partially solve the problem of vertical and front-back localisation [44], however it has to be custom-designed for every player, otherwise they will be presented with an inaccurate sound stage [44]. This would result in players not being able to locate their enemies.

Moving on to separate spatial localisation systems - distance attenuation consists of intensity and frequency loss. Point source sound intensity will drop based on real world inverse square law [38]. Frequency loss is dependent on temperature, moisture levels, atmospheric pressure, air absorption, the intensity of the wind and its direction. Also called directivity, oblique angle attenuation explains how the sound source frequency and intensity changes based on the direction it is facing, relative to the listener [4, 32].

Audio Occlusion is a reduction in loudness from passing through geometry, involving complex propagation and diffraction [33]. A diagram of occlusion can be seen in Figure 1. This is mostly simulated using low-pass filters and a reduction of intensity [5]. In addition to occlusion, portalling helps with propagating sound through an opening in the room, such as an open window or a door [6, 33]. Diffraction is the bending of sound waves around obstacles [37]. It is one of the contributors to sounds that are heard from around the corner from the emitter, for example.

Early sounds bounce off surfaces in a room and arrive between 50ms to 80ms after the direct sound reaches listeners' ears. They increase the perceived presence of the direct sound [35]. In addition, early sounds help with understanding the size of the room and proximity to nearby walls [47]. Late reflections describe the individual delays, heard after the early sounds, which are much closer together and form reverberation [35]. In-game reverberation helps with delivering the experience of being immersed in a particular virtual space [40]. In addition, standing waves provide the room with a characteristic [35]. In case of game sound - digital reverberation plugins can be used, which sometimes simulate early sounds without the need for any additional software.

### 1.2 First-Person Shooters

First-person shooters (FPS) are video games which feature the use of either realistically modelled, fantasy or science-fiction inspired guns for the purpose of killing virtual enemies, presented to the

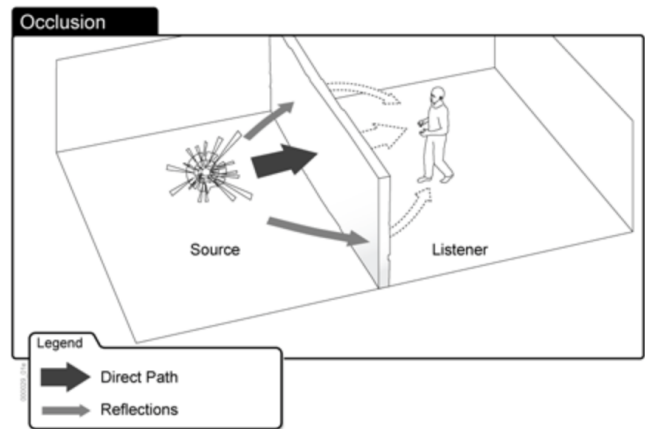


Figure 1: Occlusion diagram [5]

player in a perceived three-dimensional environment [26]. Auditory cues assist players with navigation in virtual environments [12] so it is important to design sounds that reflect and give feedback to the actions that players take. In-game character voice-over has been found to be an essential part of the game, such as a teammate calling out a grenade throw [17]. Guns have a special importance in FPS games as they are on the player's screen at all times. Poorly designed gunshot sounds can negatively impact the perception of a weapon's power, in some cases crippling a player's performance [28].

It is claimed that enemy spatial localisation plays an important role within FPS games [42]. Acousmatic sound is that which is heard without the source of the sound being visible. Acousmatic non-player character (NPC) sound design is used throughout *Silent Hill 4: The Room* [29]. By using this method, a sound designer can influence the player to perceive the NPC as a potential threat, due to the uncertainty of sound emitter's location and the NPC's intent [19]. In addition, NPCs chatting in the distance have been found to increase threat perception, as they can suggest the presence of a nearby enemy [30]. Footsteps and other sounds associated with the presence of a nearby enemy could act as a forewarning and create suspense [46]. Guillaume [46] cites fear of the unseen as an explanation of why this works. He also states that feelings of fear and discomfort in horror games are achieved with the use of auditory cues that are deliberately harder to localise and identify.

Headphone use can increase levels of immersion, as they block out sounds from the user's environment. This helps to encapsulate the player in their character, particularly in situations such as those where a character's breathing is perceived as coming from the player [25]. That said, increased quality of sound can go unnoticed by the player due to significant visual bias in an interactive environment [43]. During the course of the study, headphone use preference on various gaming platforms will be inspected to understand current use trends.

### 1.3 Threat Perception in Video Games

Threat perception is the estimation of danger [15] and depends on multiple factors - empathy towards the in-game player character

controlled by the player [31], level of immersion [13] and threat learning [1]. Players are also more likely to look out for a threat and identify it if aroused, possibly due to immersion into the game [20]. Lowest arousal levels were reported when players observed enemies on screen, while high levels were reported when players ran out of ammo [59]. Auditory figure-ground segregation can be reduced due to high visual load on the player [34]. Thus, detailed spatial auditory cues, mimicking real-world counterparts, could prove to be a useful game mechanic in attempting to identify sound sources and their positions.

Threats are communicated in multiple ways [16] such as distance attenuation, with the use of a mix of volume and balance, and timbre, which contributes towards enemy-category identification. The player's footsteps on a particular material can act as an alert to a particular enemy, occupying that specific area. The absence of an enemy visual on screen in horror games can provide the most frightening experience, leaving the interpretation of purely auditory cues up to each players' individual imagination [45]. Hearing sounds from behind while running away from an enemy was also found to increase threat [60]. A study on instilling fear with the use of auditory cue timing and volume in horror games [56] has suggested evaluating other methods of sound design, and in genres other than horror.

## 1.4 Research Question and Hypotheses

Spatial audio presence in FPS games is assumed to bring benefits. This has not, however, been academically studied [42]. The aim of the gamers is to evaluate the source of the sound produced as well as its spatial placement in relation to the player [46]. The availability of spatial composition techniques to sound designers is one factor that dictates what is being used. However, the effectiveness of specific techniques is not mentioned [7].

Ultimately the goal of this research is to answer the following question. How do different spatial audio cues affect the perception of threat in a first-person shooter game? Answering this will be important for our understanding of whether spatialised audio is a necessary tool for video game sound designers. In addition, knowing which spatial audio component provides which type of experience is useful for future in-game sound design.

In addition to threat, it was interesting to find out which system causes the enemy to be perceived as stealthier, what factors assist with position perception, and which spatial audio techniques make the audio sound realistic. Multiple hypotheses have been proposed:

- H1** – Precise identification of an opponent in relation to the player character can lower the perceived threat.

- H2** – Sonic occlusion is one of the main contributors towards an increase in perceived threat.

- H3** – Having multiple spatial audio localisation systems in-game will contribute towards higher perceived realism in FPS games.

- H4** – Occlusion can play a major role in how stealthy an enemy is perceived as being.

## 2 METHOD

This section will discuss the method of preparation for the experiment and, its design and procedure.

### 2.1 Requirements Gathering

In order to prepare the experiment, interviews with game audio professionals were carried out. Participants were asked about their experience with spatial localisation techniques, and the benefits and challenges they presented.

### 2.2 Experiment Design

**2.2.1 Questionnaire.** A listening study was suggested, the results of which would enable us to rate the performance of audio systems [9]. Multiple visual scenes needed to be designed in order to assess the efficiency of each system in different environments such as open space, closed space and mixed environments. Visual, on-screen events can increase cognitive load and bias the experiment [34]. To counter this, the enemy was hidden from the player's view. In addition, all survey pages containing audio-visual material were randomized - firstly to decrease the effect of survey fatigue [21], and also to ensure that participants are not able to predict the sequence and/or communicate it to other participants.

One of the requirements from participants was that they must be a video game player. To establish their game genre preferences and play habits, an introductory questionnaire was composed using a most commonly played game genre list from ID Tech [57]. As discussed previously, HRTF-based spatial localisation is very dependent on whether headphones are used by the players. During requirements gathering phase it was suggested that audio professionals have established a rise in PC headphone use. In order to verify this, a question regarding headphone use while playing games on different platforms was set up. Based on the hypotheses set, and the identification in the literature review of multiple systems being useful in spatial localisation, multiple questions were designed to gather information about the user experience. All four questions described below were asked after each new clip was shown to the participants.

**Q1** – “How threatening was the enemy?” was designed to address hypothesis H1 and H2.

**Q2** – “How easy was it to identify where the enemy was positioned?” was also designed to address hypothesis H1.

**Q3** – “How stealthy was the enemy?” was designed to address H4.

**Q4** – “Did the clip sound realistic?” was designed to address H3.

Questions 1, 3 and 4 used a four-point scale based on Section L of the National Center for PTSD published paper “Deployment Risk Resilience Inventory-2” [58]. This paper sets recommendations on how to measure events such as these which fits perfectly with the type of responses collected in this study. The scale includes responses such as: Not at all; A little; Moderately; and A great deal. Q2 uses a different, five-point scale, because instead of asking about the magnitude of the event, it asks about the ease of performing a task. The scale goes from ‘Very Hard’ to ‘Very easy’ with a middle step, ‘Neither hard nor easy’. In the end, two text boxes were presented to the participant so they could comment further and highlight what they liked and disliked about the sound of the clips. At the end, participants were asked about their confidence during the study, using a five-point scale, ranging from extremely confident to not at all confident.

**2.2.2 Game Scene and Audio Design.** As discussed in previous chapters, enemy footsteps, alongside talking and shooting, have been found to signify a significant threat to the player. A sequence of running, reloading, shooting (while turning away) and then running back was used as the auditory cue to alert the player to an enemy presence. All in-game sounds were recorded by the researcher apart from a gun reload sound, which was sourced from a library [11]. As discussed previously, Wwise middleware has many benefits over Unreal Engine's built-in audio engine. Therefore, Wwise was used for authoring the sounds used in the experiment. Due to a limited knowledge of 3D object design and texturing, as well as limited time, commercially available assets purchased from Unreal Marketplace [22] were used. "Environment Set" [36] was used for the forest scene. "Abandoned Factory" [48] was used for the industrial scene. "HQ Residential House" [39] was used for the indoor scene as seen in Figure 2.

Once all of the sounds were implemented, tested and calibrated for loudness [27], screen captures of the walkthroughs were made using OBS [41]. Eighteen clips were recorded, each 20 seconds in length. Six clips for each of the three scenes allowed us to test each spatial audio system separately. The first clip included only distance attenuation, which would be present in all of the other clips. The other clips added cone attenuation, reverberation and portals, reflection and occlusion – one system per clip, with the sixth and last clip having all of the above-mentioned systems enabled.

Due to the nature of the experiment and the use of a specific kit, the benefits of an offline survey app were considered. Commercial solutions, which offer offline functionality, either limit response amounts or provide randomisation as a premium feature [50, 53, 54]. In addition, they mostly offer only mobile iOS apps, rather than laptop MacOS versions. This was the reason for using a customised PHP/HTML/CSS based solution.

Due to the lockdown associated with the outbreak of SARS-CoV-2 virus (COVID-19), all face-to-face experiments and studies have been restricted by Edinburgh Napier University, following government advice. At that time, only nine responses had been gathered, and the decision was made to convert the experiment into a remote one.

To ensure volume did not exceed the experiment specifications, an iOS app called "Decibel Meter Sound Detector" [55] was used to record the average and peak volumes on the hardware used during the experiment. The values were recorded as 68 dBA average and 84 dBA peak. They were taken by placing the microphone of the iOS device to the headphone driver. Instructions were given to match the volume levels described above by increasing or decreasing volume levels on their computer. Participants were advised to switch off any applications such as operating system notifications. Android users were asked to use the "Sound Meter" app [52] instead.

This calibration was carried out using an audio clip created for this express purpose, using the loudest in-game elements – gunshot sounds. A long sequence of gunshot sounds, hosted on Soundcloud, provided the sustained loudness and peaks recorded in the experiment design. Once the calibration was complete, participants were given a link to the online version of the experiment.

Offline participants were given a participant information sheet, a privacy notice sheet and asked to sign an informed consent sheet. A MacBook [2] and a set of Sennheiser HD25 [49] headphones were

provided for them to carry out the study. All offline participants were from London, UK, as they were easier to get to.

Remote participants were sent the same participant information and asked to sign and return the informed consent sheet. They were given a link to download one of two sound meter apps for their phones and asked to match the loudness level. Once this process was completed, they were given an online link to perform the study. Remote participants were based all around the world.

## 3 RESULTS

The results of the studies will be discussed in two separate chapters. The first will look at the interviews conducted during the requirements gathering stage. The following chapter will examine the experiment results.

### 3.1 Requirements Gathering

Two relevant game audio professionals agreed to an interview for the requirements gathering stage. Both participants were male, aged 36-47 years, with an average industry experience of 13 years. One is currently an audio director, while the other is a senior sound designer. Both are located in Europe.

The interviewees have both found spatial localisation to be very important in online FPS games, due to players' need for precise position identification in order to eliminate their enemy. Games that use hard panning, with sound only playing in one ear, were highlighted by the participants as bad examples as they tended to sound very unnatural. Both participants mentioned this issue. In addition, HRTFs were mentioned as another effective method of spatial localisation, even though it needs more work to provide a transferable experience to all gamers. HRTFs, in addition to room acoustics, occlusion and obstruction simulations, were noted as primary contributors towards the identification of enemy position. One interviewee stated that the absence of a precise spatial localisation system within a game would increase his perception of threat, due to the subsequent inability to accurately pinpoint an enemy's location.

Both participants identified occlusion as an essential technique. It was noted that simple, filter-based systems do not sound realistic and should be avoided entirely if a complex, more life-like system is not accessible by the audio designer. In-game audio physics simulation has to be as precise as possible to mimic the real world, as that reduces the player's in-game learning curve.

Methods of authoring audio attenuation over distance included the use of content recorded at various distances from the source and switching or crossfading between them in-game. First-order reflections were either handled by pre-recording sounds, or with the use of raycasting with subsequent application of time delay based on the result. It was noted that smaller studios have to simulate first-order reflections virtually, due to budget constraints which do not allow for separate recordings of every gun in the field.

### 3.2 Experiment

In total, 35 participants (7 females, 20%) took part in the experiment. Of these, 9 took the offline version. Remote participant replies were not too different to offline ones. All had normal hearing and normal or corrected eyesight. Participants were aged between 20 and 44





Figure 2: From left to right: house scene, industrial scene and forest scene

years old (mean 29.5). All participants reported playing video games regularly. As can be seen in Figure 3, the most played genre was adventure, chosen by 32 (91%) participants. Role-playing games took second place with 29 (83%). First-person and third-person shooters took third place with 27 (77%) each. Sports was the least played genre with only 18 people (51%).

Headphone use and gaming platform preference, were next. Results in Table 1 show the percentage of participants who do not use, or own, a specified gaming platform. Separate counts for each platform are included below that. We can see that the least popular platforms are VR and AR, with just 23% of people using them. The most popular was Mobile (86%), followed by Desktop PC (83%). Consoles came third, used by 80% of participants.

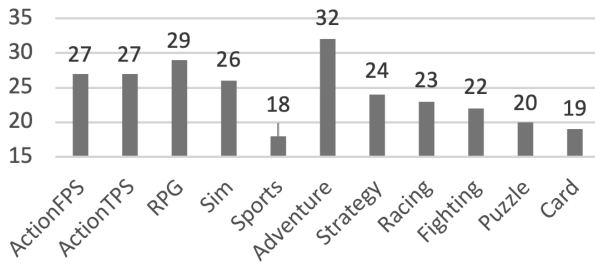


Figure 3: Participant played game genres

It seems that desktop and laptop gamers mostly prefer to use headphones during gameplay. The majority (73%) use headphones all or most of the time while playing on a desktop PC. The opposite is true for console gamers – 75% of whom do not, or almost never, use headphones while playing games.

Table 2 shows the mean results from all scenes. All of the tested systems are displayed on the left, while the measured parameters are shown on the top. Apart from the position, which used a five-point scale, all of the other parameters used a four-point scale. The higher the number, the more effective the spatial localisation system is at achieving the measured parameter.

Line charts have been used as a suggested method to visualise and then analyse the data [8]. Using line charts, the table has been marked up with highest values per parameter in bold, and the lowest in italics.

	Desk top	Lap top	Con sole	Port able	Mob ile	VR	AR
Always	10	10	1	3	4	6	1
Mostly	11	10	3	3	11	2	0
Sometimes	7	7	3	7	7	0	2
Almost Never	1	0	7	5	5	0	2
Never	0	0	14	6	3	0	3
Do Not Use	6	8	7	11	5	27	27

Table 1: Headphone use on various platforms

Mean replies for all scenes combined and individually, did not score less than 2 or more than 3 in the threat parameter (apart from just one exception – the Forest Full scene at 1.97). This places most systems in the “a little” to “moderately” effect category. It is possible that participants could have preferred these options to “not at all” as the enemy was invisible.

The interesting thing here is that occlusion is not the main contributor towards a player’s assessment of an enemy position. As we see from the Position tab in Table 2, occlusion and full clips scored below 3, while all others scored more than 3.1. Standard deviation is included after each median figure.

	Threat	Position	Stealth	Realism
Full	2.23/1.02	2.48/1.16	<b>2.49/0.98</b>	2.65/1.04
Basic	<b>2.57/0.86</b>	<b>3.19/1.05</b>	1.93/0.97	2.49/0.94
Reverb/Portal	<b>2.59/0.96</b>	<b>3.10/1.08</b>	2.09/0.86	<b>2.81/0.83</b>
Reflectors	<b>2.54/0.82</b>	<b>3.17/1.05</b>	2.05/0.86	2.69/0.85
Cone Attn.	2.35/0.85	<b>3.18/0.99</b>	2.00/0.82	2.49/0.87
Occlusion	2.10/0.76	2.92/1.14	<b>2.45/0.92</b>	2.45/0.96

Table 2: Mean results from all scenes

Many more systems have been effective in helping participants assess enemy location. Reverberation and portalling, material reflectors and cone attenuation were all graded highly. Surprisingly, the basic clip scored highest of all. This suggests it is as effective as the aforementioned systems in aiding enemy character spatial localisation. The full-system implementation, meanwhile, scored lowest of all.

Both occlusion and the full-system implementation scored highest on enemy stealth. The full system had occlusion turned on as well, so this could be a factor in the result. The basic system scored the lowest. Finally, realism was most positively influenced by the use of reverberation and portalling. The results show us that cone

attenuation, occlusion and basic systems all contributed towards less realistic outcomes.

All of the results for the forest scene are provided in Table 3. As with the previous table, all of the lowest and highest values are highlighted, following a line chart analysis. Standard deviation is included after each median figure.

	Threat	Position	Stealth	Realism
Full (F)	1.97/1.06	2.34/1.16	<b>2.54</b> /1.01	2.31/1.10
Basic (F)	<b>2.63</b> /0.91	<b>3.03</b> /1.12	1.97/0.98	2.37/0.94
Reverb/Portal (F)	2.26/0.92	2.66/1.21	2.11/0.80	<b>2.57</b> /0.78
Reflectors (F)	2.40/0.77	<b>3.03</b> /1.18	2.00/0.87	2.43/0.85
Cone Attn. (F)	2.20/0.72	2.89/0.93	2.20/0.90	2.31/0.72
Occlusion (F)	2.00/0.73	2.66/1.24	<b>2.43</b> /1.04	2.14/0.91

**Table 3: Mean results from the Forest scene (F)**

As we can see from the results, the biggest impact on threat perception was achieved by the basic clip. Occlusion and the full implementation seem to have a negative effect on threat perception in this scene. Similarly, the full clip has also had a negative impact on the perception of enemy position, along with the basic and the reflector-aiding clips. Once again, stealth is observed to have been best aided by occlusion and the full clip, which contains occlusion. Reverberation was the primary contributor towards realism within the scene, with occlusion contributing negatively.

In a scene with both an open environment and buildings close to the player, reverb and material reflections seem to have made the enemy sound more threatening, as shown in Table 4. Occlusion, once again, appears to decrease threat. Interestingly enough, even though most components have scored highly in helping to identify enemy position, the full clip, containing all of them sees a significant drop in the rating. It is also worth noticing that occlusion is not a contributor towards position identification in the Industrial scene.

	Threat	Position	Stealth	Realism
Full (I)	2.51/1.09	2.43/1.14	<b>2.54</b> /1.01	<b>3.03</b> /0.98
Basic (I)	2.34/0.91	<b>3.51</b> /1.01	1.69/0.90	2.37/1.03
Reverb/Portal (I)	<b>2.91</b> /0.92	<b>3.37</b> /1.03	2.03/0.92	<b>3.11</b> /0.83
Reflectors (I)	<b>2.74</b> /0.89	<b>3.40</b> /1.06	1.97/0.86	2.80/0.83
Cone Attn. (I)	2.31/0.93	<b>3.40</b> /1.01	1.74/0.70	2.43/0.92
Occlusion (I)	2.23/0.73	3.09/1.04	<b>2.31</b> /0.80	2.51/0.95

**Table 4: Mean results from the Industrial scene (I)**

As with previous results, occlusion is found to contribute towards stealth, along with the full clip. The basic and cone attenuation clips appear to negatively impact stealth. Reverberation is also shown to increase perceived realism within the game, as does the full clip.

Finally, as we can see from Table 5, occlusion seems to bring down threat levels in the house scene. The basic clip once again increases threat perception in participants. Reverberation and cone attenuation were main contributors towards ease of position identification, with the full implementation scoring lowest in this category. Again, stealth was mostly affected by occlusion, while cone attenuation appears to be the main contributor towards decreased

stealth. For this scene, reflectors were identified as a system, making the scene sound realistic, even if within a minimal margin.

	Threat	Position	Stealth	Realism
Full (H)	2.20/0.87	2.66/1.21	2.37/0.94	2.60/0.95
Basic (H)	<b>2.74</b> /0.74	3.03/0.95	2.14/1.00	2.71/0.83
Reverb/Portal (H)	2.60/0.95	<b>3.29</b> /0.86	2.11/0.87	2.74/0.82
Reflectors (H)	2.49/0.78	3.09/0.89	2.17/0.86	<b>2.83</b> /0.82
Cone Attn. (H)	2.54/0.89	<b>3.26</b> /0.98	2.06/0.80	2.71/0.93
Occlusion (H)	2.09/0.82	3.03/1.12	<b>2.60</b> /0.91	2.69/0.96

**Table 5: Mean results from the House scene (H)**

Quite a few comments mentioned feeling more threat when participants thought the enemy was very close to them. The results have shown that most reported this occurring when there was almost no spatialisation system operating. Participants also mentioned that they found some enemy sounds unconvincing. They reported that the sounds seemed to come from the player. This is a direct by-product of an absent spatial localisation system, which was already established as a negative contributor to realism.

Participants were also asked to explain what they felt would help them best with enemy localisation in an FPS game. Many people described the effects of occlusion, noting that it should be a consistent experience when objects are obstructing a player's view of the enemy. Some felt that more aggressive panning would help them understand more clearly where the enemy is located. Individual objects equipped with their own auditory cues were also mentioned, as opposed to a single stereo ambience track containing all objects in the scene. It was felt that having a reference to other objects indoors would help as a point of comparison to the sound of the enemy, for instance to better understand how far away the enemy might be. Shooting sounds were also noted as proving helpful in localisation.

## 4 DISCUSSION

Within this study different spatial localisation methods affected the perception of threat, stealth, realism and perception of enemy location in different ways. In an open scene, with sparsely placed objects, occlusion can have a negative effect on threat perception, while a basic left-right panning mechanism, combined with distance attenuation can measurably increase it.

For stealthy action, occlusion is the best tool. However, if stealth is to be avoided then the basic system, potentially with added material reflections, should suffice. Finally, occlusion could drive down the perceived realism of sound in an open space scene. In order to achieve a more realistic scene – reverberation and portalling is advised.

For a scene in which a player is surrounded by objects, such as a city, deploying reverberation and material reflectors could help to increase the perception of enemy threat. Reverberation can also be used to make the scene sound more realistic.

Within closed environments such as a house, having no spatialisation systems in place could make players feel enemy threat more acutely. Adding reverberation helps with location perception,

as does cone attenuation. However, this gain in spatial localisation could come at a cost to stealth perception. Early reflections or material reflectors could make the scene more realistic.

Following the research, hypothesis H1 was supported by the findings. The ability of players to locate an enemy using spatial audio lowers the perceived threat of that enemy. It seems that the inability to place the enemy using auditory cues combined with the absence of the enemy on screen, has the effect of increasing perceived threat.

Presence of occlusion in a game appears to lower the perceived threat, as the enemy can be located easily as well as identified as being behind a boundary, away from the player. These findings address the hypothesis H2. The absence of a direct view of the enemy, along with auditory cue confirmation of the enemy being hidden away appears to reduce feelings of impending danger in the player. While providing additional spatial information to the player, it does not increase perception of threat.

The third hypothesis is supported with one exception. Having multiple in-game spatial audio localisation systems will contribute towards higher perceived realism, except in the case of occlusion. Occlusion, either due to the way it is implemented in Wwise, or due to filtering of the source sound, was not quite as effective throughout the experiment. Reverberation, on the other hand, has been identified as one of the main contributors to audio realism.

Finally, addressing the last hypothesis – occlusion does indeed play a significant role in stealth perception. Compared to other spatial audio techniques, it seems to be the most consistent – indeed only – factor in all scenes.

#### 4.1 Contribution

Sound designers and technical sound designers can plan their scenes with the use of information gathered in this experiment. Realistic scenes would have to have reverberation, unless they are based outdoors. On the other hand, experiments could be carried out for scenes that are intended to convey an unrealistic scenario, perhaps a dream sequence or a scene that is imagined by a character. Omitting reverberation within that scene could suggest to players that this is not a realistic world and enhance the story sonically.

When building a stealth-based game, occlusion should, ideally, be used in the scene. It is the primary contributor to stealth and is essential to paint the enemy as such. Occlusion absence could be used to convey an enemy who is actively trying to be noticed, however, the absence of occlusion could also have a negative effect on the identification of enemy position.

Indoor scenes would benefit from early reflections a lot more than any other. Audio designers could put more resources into recorded or virtual simulations of those spaces with regard to early reflections in order to help with the player-perceived degree of realism. Having only distance attenuation in an indoor scene was seen to have a big impact on threat perception. If a player is required to feel precisely that, then audio designers could use this tactic to get that outcome.

#### 4.2 Future Work

While the results were reasonably consistent, even considering different methods of experiment delivery, a more consistent approach

could be used. In addition, a study of a similar nature could be performed on a more significant number of people. Instead of targeting all gamers, the experiment could have used only FPS players to see what works best for them and whether their previous experience might have an effect on outcomes. It would be beneficial to differentiate players who use headphones for FPS games exclusively, rather than all games on PC, to see if the number of users stays the same or whether the percentage rises.

Position perception has been evaluated by the participants in this experiment, however, we do not know how accurate their localisation was. An experiment to test this could be devised with different spatialisation systems, where players point at a place on the game map to indicate where they thought the enemy was. Q4 was found to have not been worded correctly to determine whether the auditory experience was realistic. Participants commented on an individual sound source, rather than the overall experience of the scene. Due to this research work being a first iteration, any future experiments should use "believable" instead of "realistic".

While the experiment has been focused on nearby enemies, it would be worth investigating spatial localisation in scenes of a more significant size, for instance those in the Battle Royale genre. As this is a relatively new arrival, there is little available research. The open nature of the game level could require a different design method.

#### ACKNOWLEDGMENTS

I would like to thank all of the participants who took part in the studies.

#### REFERENCES

- [1] Annemieke M. Apergis-Schoute, Daniela Schiller, Joseph E. Ledoux, and Elizabeth A. Phelps. 2014. Extinction resistant changes in the human auditory association cortex following threat learning. *Neurobiology of Learning and Memory* 113 (2014), 109–114. <https://doi.org/10.1016/j.nlm.2014.01.016>
- [2] Apple. 2011. MacBook Pro. <https://www.apple.com/uk/macbook-pro/>
- [3] Audiokinetic. 2019. Wwise.
- [4] Audiokinetic. 2020. Adding Cone Attenuation. [https://www.audiokinetic.com/courses/wwise101/?source=wwise101&id=adding\[\\_\]cone\[\\_\]attenuation](https://www.audiokinetic.com/courses/wwise101/?source=wwise101&id=adding[_]cone[_]attenuation)
- [5] Audiokinetic. 2020. Obstruction and Occlusion in Environments. [https://www.audiokinetic.com/library/2018.1.11\[\\_\]6987/?source=SDK&id=soundengine\[\\_\]jobsoc.html](https://www.audiokinetic.com/library/2018.1.11[_]6987/?source=SDK&id=soundengine[_]jobsoc.html)
- [6] Audiokinetic. 2020. Using Wwise Spatial Audio in Unreal. [https://www.audiokinetic.com/library/edge/?source=UE4&id=usingwwisespatialaudio.html\[#\]sectionC](https://www.audiokinetic.com/library/edge/?source=UE4&id=usingwwisespatialaudio.html[#]sectionC)
- [7] Marije A.J. Baalman. 2010. Spatial composition techniques and sound spatialisation technologies. *Organised Sound* 15, 3 (2010), 209–218. <https://doi.org/10.1017/S1355771810000245>
- [8] Juhee Bae, Göran Falkman, Tove Helldin, and Maria Riveiro. 2019. Visual Data Analysis. In *Data Science in Practice*. Chapter 8, 133–155. [https://doi.org/10.1007/978-3-319-97556-6\\_8](https://doi.org/10.1007/978-3-319-97556-6_8)
- [9] Søren. Bech and Nick. Zacharov. 2007. *Perceptual audio evaluation : theory, method and application*. John Wiley & Sons. 442 pages.
- [10] Durand R. Begault and Elizabeth M. Wenzel. 1993. Headphone Localization of Speech. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 35, 2 (jun 1993), 361–376. <https://doi.org/10.1177/001872089303500210>
- [11] BOOM Library. 2019. Gun Sounds. <https://www.boomlibrary.com/sound-effects/gun-sounds/>
- [12] James Broderick, Jim Duggan, and Sam Redfern. 2018. The Importance of Spatial Audio in Modern Games and Virtual Environments. In *2018 IEEE Games, Entertainment, Media Conference, GEM 2018*. IEEE, 507–510. <https://doi.org/10.1109/GEM.2018.8516445>
- [13] Christopher G. Burns and Stephen H. Fairclough. 2014. Use of auditory event-related potentials to measure immersion during a computer game. *International Journal of Human Computer Studies* 73 (2014), 107–114. <https://doi.org/10.1016/j.ijhcs.2014.09.002>

- [14] Robert A Butler and Richard A Humanski. 1992. Localization of sound in the vertical plane with and without high-frequency spectral cues. *Perception & Psychophysics* 51 (1992), 182–186. <https://doi.org/10.3758/BF03212242>
- [15] Raymond Cohen. 1978. Threat Perception in International Crisis. *Political Science Quarterly* 93, 1 (1978), 93. <https://doi.org/10.2307/2149052>
- [16] Flávio Coutinho, Raquel O. Prates, and Luiz Chaimowicz. 2011. An analysis of information conveyed through audio in an FPS game and its impact on deaf players experience. *Brazilian Symposium on Games and Digital Entertainment, SBGAMES step 1* (2011), 53–62. <https://doi.org/10.1109/SBGAMES.2011.16>
- [17] Sebastian Darth. 2015. "How does Sound Design Impact the Players Perception of the Level of Difficulty in a First Person Shooter Game?". Technical Report. Luleå University of Technology. 45 pages.
- [18] Dice. 2018. Battlefield V Dev Talks: The Audio of Battlefield V - YouTube. <https://www.youtube.com/watch?v=i-6OxRBHG3o>
- [19] K Donnelly. 2016. Emotional Sound Effects and Metal Machine Music: Sound-worlds in Silent Hill Games and Films. In *The Palgrave Handbook of Sound Design and Music in Screen Media*, L Greene and D Kulezic-Wilson (Eds.). Chapter 6, 73–88. [https://doi.org/10.1057/978-1-137-51680-0\\_6](https://doi.org/10.1057/978-1-137-51680-0_6)
- [20] Robert Dudley, Guy Dodgson, Gabrielle Sarll, Racheal Halhead, Helen Bolas, and Simon McCarthy-Jones. 2014. The effect of arousal on auditory threat detection and the relationship to auditory hallucinations. *Journal of Behavior Therapy and Experimental Psychiatry* 45 (2014), 311–318. <https://doi.org/10.1016/j.jbtep.2014.02.002>
- [21] Brian L. Egleston, Suzanne M. Miller, and Neal J. Meropol. 2011. The impact of misclassification due to survey response fatigue on estimation and identifiability of treatment effects. *Statistics in Medicine* 30, 30 (dec 2011), 3560–3572. <https://doi.org/10.1002/sim.4377>
- [22] Epic Games. 2020. UE4 Marketplace. <https://www.unrealengine.com/marketplace/en-US/store>
- [23] Epic Games. 2020. Unreal Engine. <https://www.unrealengine.com/en-US/>
- [24] Mark Grimshaw. 2007. The Resonating spaces of first-person shooter games. In ... of The 5th International Conference on Game .... [http://sprite.bolton.ac.uk/2411/gcct/\\_conferencepr-4.pdf](http://sprite.bolton.ac.uk/2411/gcct/_conferencepr-4.pdf)
- [25] Mark Grimshaw. 2008. Sound and immersion in the first-person shooter. *UBIR: University of Bolton Institutional Repository* (2008). [http://digitalcommons.bolton.ac.uk/gcct/\\_journalspr/3](http://digitalcommons.bolton.ac.uk/gcct/_journalspr/3)
- [26] Michael Hitchens. 2011. Game Studies - A Survey of First-person Shooters and their Avatars. [http://gamestudies.org/1103/articles/michael\\_hitchens](http://gamestudies.org/1103/articles/michael_hitchens)
- [27] HM Government. 2005. The Control of Noise at Work Regulations.
- [28] Zander Hulme. 2018. Wolfenstein Game Audio Feel. <https://clips.twitch.tv/SpeedyClumsySlothUWot>
- [29] Konami. 2004. Silent Hill 4: The Room.
- [30] Daniel Kromand. 2008. Sound and the diegesis in survival-horror games. In *Proceedings of the Audio Mostly Conference - A Conference on Interaction with Sound*, 16–19.
- [31] Petri Lankoski. 2011. Player Character Engagement in Computer Games. *Games and Culture* 6 (2011), 291–311. <https://doi.org/10.1177/1555412010391088>
- [32] Mc Squared System Design Group. 2020. What is Directivity? <http://www.mcsquared.com/directvt.htm>
- [33] Microsoft. 2017. Project Triton - Microsoft Research. <https://www.microsoft.com/en-us/research/project/project-triton/>
- [34] Katharine Molloy, Nilli Lavie, and Maria Chait. 2019. Auditory figure-ground segregation is impaired by high visual load. *Journal of Neuroscience* 39, 9 (2019), 1699–1708. <https://doi.org/10.1523/JNEUROSCI.2518-18.2018>
- [35] David Murphy and Flaithri Neff. 2010. Spatial sound for computer games and virtual reality. In *Game Sound Technology and Player Interaction: Concepts and Developments*, 287–312. <https://doi.org/10.4018/978-1-61692-828-5.ch014>
- [36] NatureManufacture. 2017. Environment Set. <https://www.unrealengine.com/marketplace/en-US/product/environment-set>
- [37] R Nave. 2009. Diffraction of Sound. <http://hyperphysics.phy-astr.gsu.edu/hbase/Sound/diffract.html>
- [38] R Nave. 2016. Inverse Square Law for Sound. <http://hyperphysics.phy-astr.gsu.edu/hbase/Acoustic/invsqs.html#c1>
- [39] NOTLonely. 2016. HQ Residential House. <https://www.unrealengine.com/marketplace/en-US/product/hq-residential-house>
- [40] Oscar Nyberg. 2019. Does changing the frequency content of a reverb affect perceived emotion in a video-game? Technical Report. Luleå University of Technology. 36 pages.
- [41] OBS. 2020. Open Broadcaster Software. <https://obsproject.com/>
- [42] Joe Rees-Jones, Jude Brereton, and Damian Murphy. 2015. Spatial audio quality and user preference of listening systems in video games. In *DAFx 2015 - Proceedings of the 18th International Conference on Digital Audio Effects*, 1–8.
- [43] Joe Rees-Jones and Damian Murphy. 2017. Spatial quality and user preference of headphone based multichannel audio rendering systems for video games: A pilot study. *142nd Audio Engineering Society International Convention 2017, AES 2017* (2017).
- [44] M Risoud, J.-N Hanson, F Gauvrit, C Renard, P.-E Lemesre, N.-X Bonne, and C Vincent. 2018. Sound source localization. *European Annals of Otorhinolaryngology, Head and Neck diseases* 135 (2018), 259–264. <https://doi.org/10.1016/j.anorl.2018.04.009>
- [45] Rebecca Roberts. 2014. Fear of the unknown music and sound design in psychological horror games. In *Music in Video Games: Studying Play*, K Donnelly (Ed.). Routledge, Chapter 7, 138–150. <https://doi.org/10.4324/9781315882697>
- [46] Guillaume Roux-Girard. 2010. Listening to fear: A study of sound in horror computer games. In *Game Sound Technology and Player Interaction: Concepts and Developments*, IGI Global, Chapter 10, 192–212. <https://doi.org/10.4018/978-1-61692-828-5.ch010>
- [47] Egil Sandfeld. 2019. Acoustic Simulation Using Wwise. <https://blog.audiokinetic.com/acoustic-simulation-with-wwise/>
- [48] Scans Factory. 2019. Abandoned Factory Buildings - Day/Night Scene. <https://www.unrealengine.com/marketplace/en-US/product/scans-abandoned-factory-buildings-day-night-scene>
- [49] Sennheiser. 2017. HD 25. <https://en-uk.sennheiser.com/on-ear-dj-headphone-hd25>
- [50] SmartSurvey. 2020. SmartSurvey. <https://www.smartsurvey.co.uk/>
- [51] Fabio Wanderley Janhan Sousa. 2011. Subjective comparison between stereo and binaural processing from B-format Ambisonic raw audio material. In *130th Audio Engineering Society Convention 2011*, Vol. 1, 548–554.
- [52] Splend Apps. 2020. Sound Meter. <https://play.google.com/store/apps/details?id=com.splendapps.decibel&hl=en>
- [53] SurveyMonkey. 2020. SurveyMonkey. <https://www.surveymonkey.com/>
- [54] TabbleDabble. 2020. Quick Tap Survey. <https://www.quicktapsurvey.com/>
- [55] Ashraf Thoppukadavil. 2020. Decibel Meter Sound Detector. <https://apps.apple.com/gb/app/decibel-meter-sound-detector/id1254994873>
- [56] Paul Toprac and Ahmed Abdel-Meguid. 2010. Causing fear, suspense, and anxiety using sound design in computer games. In *Game Sound Technology and Player Interaction: Concepts and Developments*, IGI Global, 176–191. <https://doi.org/10.4018/978-1-61692-828-5.ch009>
- [57] Vince. 2018. Ultimate List of Different Types of Video Games | 49 Genres & Subcategories. <https://www.idtech.com/blog/different-types-of-video-game-genres>
- [58] D. Vogt, B. N. Smith, D. W. King, and L King. 2012. *Deployment Risk & Resilience Inventory-2 (DRRI-2)*. Technical Report. National Center for PTSD, Boston.
- [59] René Weber, Katharina-Maria Behr, Ron Tamborini, Ute Ritterfeld, and Klaus Mathiak. 2009. What Do We Really Know About First-Person-Shooter Games? An Event-Related, High-Resolution Content Analysis. *Journal of Computer-Mediated Communication* 14 (2009), 1016–1037. <https://doi.org/10.1111/j.1083-6101.2009.01479.x>
- [60] Joel Windels. 2011. Scary Game Findings: A Study Of Horror Games And Their Players. [https://www.gamasutra.com/view/feature/134848/scary\\_game\\_findings\\_a\\_study\\_of\\_horror\\_games\\_and\\_their\\_players.php](https://www.gamasutra.com/view/feature/134848/scary_game_findings_a_study_of_horror_games_and_their_players.php)
- [61] Shu Nung Yao. 2017. Headphone-based immersive audio for virtual reality headsets. *IEEE Transactions on Consumer Electronics* 63, 3 (2017), 300–308. <https://doi.org/10.1109/TCE.2017.014951>